

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5091122号  
(P5091122)

(45) 発行日 平成24年12月5日 (2012. 12. 5)

(24) 登録日 平成24年9月21日 (2012. 9. 21)

(51) Int. Cl.

F I

G 0 6 F 17/18 (2006. 01)

G 0 6 F 17/18

D

C 1 2 Q 1/68 (2006. 01)

C 1 2 Q 1/68

A

C 1 2 M 1/00 (2006. 01)

C 1 2 M 1/00

A

請求項の数 10 (全 13 頁)

(21) 出願番号 特願2008-511434 (P2008-511434)  
 (86) (22) 出願日 平成18年5月12日 (2006. 5. 12)  
 (65) 公表番号 特表2008-546048 (P2008-546048A)  
 (43) 公表日 平成20年12月18日 (2008. 12. 18)  
 (86) 国際出願番号 PCT/US2006/018549  
 (87) 国際公開番号 W02006/124673  
 (87) 国際公開日 平成18年11月23日 (2006. 11. 23)  
 審査請求日 平成21年3月17日 (2009. 3. 17)  
 (31) 優先権主張番号 60/681, 182  
 (32) 優先日 平成17年5月13日 (2005. 5. 13)  
 (33) 優先権主張国 米国 (US)  
 (31) 優先権主張番号 11/432, 856  
 (32) 優先日 平成18年5月11日 (2006. 5. 11)  
 (33) 優先権主張国 米国 (US)

(73) 特許権者 591099809  
 バイオーラッド ラボラトリーズ, インコ  
 ーポレイティド  
 アメリカ合衆国, カリフォルニア 945  
 47, ハーキュルズ, アルフレッド ノー  
 ベル ドライブ 1000  
 (74) 代理人 100099759  
 弁理士 青木 篤  
 (74) 代理人 100077517  
 弁理士 石田 敬  
 (74) 代理人 100087871  
 弁理士 福本 積  
 (74) 代理人 100087413  
 弁理士 古賀 哲次

最終頁に続く

(54) 【発明の名称】 統計的線形データの同定

(57) 【特許請求の範囲】

【請求項 1】

データセットが統計的線形挙動を示すかを決定するために当該データセットをコンピューターにおいて処理する方法であって、ここで当該コンピューターが、データ受け取り手段及びプロセッサを含み、

当該データ受け取り手段が、シグナル成分とノイズ成分を有するオリジナルセットのデータポイントを受け取り；

当該プロセッサが、以下の工程：

上記オリジナルセットのデータポイントを線形関数に近似し；

上記オリジナルセットのデータポイントと上記近似された線形関数との間の残差を計算し；

上記オリジナルセットのデータポイントと上記近似された線形関数との間の残差の第一尺度を計算し；

上記オリジナルセットのデータポイントに存在する上記ノイズ成分を、以下の：

i) 平滑化データポイントのセットを測定することにより平滑化データセットを計算し、ここで平滑化データポイントの数値は関数 G により測定され、そして当該平滑化データポイントの近くに位置する複数のオリジナルデータポイントの数値に基づいており、ここで当該平滑化データポイントの近くに位置するオリジナルデータポイントが、平滑化データポイントの付近のウィンドウ内に存在し；そして

i i) 上記平滑化データセットと上記オリジナルセットのデータポイントとの間の

10

20

残差を計算する

により見積もり；

上記平滑化データセットと上記オリジナルセットのデータポイントとの間の残差の第二尺度を計算し；そして

上記第一尺度と上記第二尺度を比較して上記オリジナルセットのデータポイントが統計的線形挙動を示すかを決定する

を含む、前記方法。

【請求項 2】

前記オリジナルセットのデータポイントの近似が、最小二乗法近似を使用することを含む、請求項 1 に記載の方法。

10

【請求項 3】

前記平滑化データセットの計算が、ローパスフィルターを使用することを含む、請求項 1 に記載の方法。

【請求項 4】

平滑化データポイントの数値が、当該平滑化データポイントの付近の上記ウィンドウ内に存在する上記オリジナルデータポイントの平均である、請求項 1 に記載の方法。

【請求項 5】

前記第一尺度と前記第二尺度が各々標準偏差である、請求項 1 に記載の方法。

【請求項 6】

前記比較が、前記第一尺度と前記第二尺度の比率を計算して、当該比率が所定の値より小さいか又は大きいかを決定することを含む、請求項 1 に記載の方法。

20

【請求項 7】

前記所定の値が 1 のオーダーである、請求項 6 に記載の方法。

【請求項 8】

前記データが P C R 増幅曲線を表す、請求項 1 に記載の方法。

【請求項 9】

データ処理を実行して、当該データが線形挙動を示すかを決定する命令を実行するようにコンピューターのデータ受け取り手段及びプロセッサを制御するように適用された複数の指示を格納するコンピューターが読み取り可能な貯蔵媒体であって、

当該データ受け取り手段が、シグナル成分とノイズ成分を有するオリジナルセットのデータポイントを受け取り；

30

当該プロセッサが、以下の工程：

上記オリジナルセットのデータポイントを線形関数に近似し；

上記オリジナルセットのデータポイントと上記近似された線形関数との間の残差を計算し；

上記オリジナルセットのデータポイントと上記近似された線形関数との間の残差の第一尺度を計算し；

上記オリジナルセットのデータポイントに存在する上記ノイズ成分を、以下の：

i) 平滑化データポイントのセットを決定することにより平滑化データセットを計算し、ここで平滑化データポイントの数値は、関数 G により測定され、そして当該平滑化データポイントの近くに位置する複数のオリジナルデータポイントの値に基づいており、  
ここで近くに位置するオリジナルデータポイントは、平滑化データポイントの付近のウィンドウ内に存在し；そして

40

i i) 上記平滑化データセットと上記オリジナルデータセットとの間の残差を計算する

により見積もり

上記平滑化データセットと上記オリジナルセットのデータポイントとの間の残差の第二尺度を計算し；そして

上記第一尺度と上記第二尺度とを比較して当該オリジナルセットのデータポイントが統計的に線形挙動を示すかを決定する

50

を実行する、前記媒体。

【請求項 10】

以下の：

シグナル成分とノイズ成分を有するオリジナルセットのデータポイントを産生する検出器；並びに

当該データが線形挙動を示すかを決定するためにデータを処理するプロセッサであって、以下の：

上記オリジナルセットのデータポイントを線形関数に近似し；

上記オリジナルセットのデータポイントと上記近似された線形関数との間の残差を計算し；

上記オリジナルセットのデータポイントと上記近似された線形関数との間の残差の第一尺度を計算し；

上記オリジナルセットのデータポイントに存在する上記ノイズ成分を、以下の：

i) 平滑化データポイントのセットを決定することにより平滑化データセットを計算し、ここで平滑化データポイントの数値は、関数 G により測定し、そして当該平滑化データポイントの近くに位置する複数のオリジナルデータポイントの値に基づいており、ここで近くに位置する当該オリジナルデータポイントが、平滑化データポイントの付近のウィンドウ内に存在し、当該ウィンドウが、オリジナルデータポイントのセットの全体の規模よりも小さい；そして

i i) 上記平滑化データセットと上記オリジナルセットのデータポイントとの間の残差を計算する

により見積もり；

上記平滑化データセットと上記オリジナルセットのデータポイントとの間の残差の第二尺度を計算し；そして

上記第一尺度と上記第二尺度とを比較して上記オリジナルセットのデータポイントが統計的に線形挙動を示すかを決定する

ことによるロジック

を含む P C R 検出装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、一般的にデータ処理システム及び方法、並びにより具体的にポリメラーゼ連鎖反応 (P C R) などの増幅過程において統計的線形データを同定するシステム及び方法に関する。

【0002】

関連出願へのクロスリファレンス

本出願は、2005年5月13日に提出された「SYSTEMS AND METHODS FOR IDENTIFYING STATISTICALLY FLAT DATA IN A DATA SET」と表題を付された米国仮特許出願60/681,182号の利益を主張し、当該出願は「DETERMINATION OF BASELINE END CYCLE IN REAL TIME PCR USING DERIVATIVE PEAK ANALYSIS」と表題を付された米国仮特許出願第60/680,765号(出願人整理番号第002558-072200US)と同時に提出されており、関連している。これらの文献の開示は、その全てを本明細書に援用する。本出願はまた、同一出願人による「BASELINING AMPLIFICATION DATA」と表題を付された米国特許出願第\_\_\_\_号(出願人整理番号第002558-072210US)に関連する。これらの出願の各開示は、全ての目的のため本明細書に援用される。

【背景技術】

【0003】

多くの実験的プロセスは量の増幅を示す。例えば、P C Rでは、量は複製されたDNA鎖の部分の数に相当し、この数は増幅段階又は増幅領域で劇的に増幅される。増幅を示す他の実験的プロセスは、細菌増殖過程を含む。量は、データシグナルを介して実験装置か

10

20

30

40

50

ら検出され、データシグナルのデータポイントは、増幅についての情報を決定するために分析される。データ分析の一環として、増幅が潜在的に生じているかを知ることが重要であり；このことを知らなければ、増幅が生じないデータを分析するという労力が無駄になされう。データが統計的に線形であれば、増幅は生じていない。

【0004】

理想的には、増幅検出装置から得たデータは、単調かつ連続的なシグナルであり、そうして当該データ、又はデータの一部が統計的に線形挙動を有するかを容易に同定することができる。しかしながら、増幅装置から得たシグナルは一般的にノイズを含み、そうしてシグナルの挙動を同定することが難しくなる。ノイズは、装置から得たシグナルの各データ点において現れ、真のシグナル、例えば実際のDNA鎖の数、の上に生じるランダムな揺らぎを与える。こうして、線形挙動を同定するためには、データを処理することが必要となる。

10

【0005】

統計的に線形であるかを決定するための一般的な従来のデータ処理方法は、線形最小二乗法(LSQ)近似を用いる。LSQ近似の相関値は、適切な近似が存在するかを決定するために使用することができる。標準的な変換により、0の相関値は、近似しないこと(bad fit)に関し、そうして当該データは線形ではなく、そして1の相関値は線形について良好な近似を示唆する。ノイズの存在下では、統計的に線形に見えるデータについて相関値が0又は1の近くになってしまうことがあるという問題点がある。さらに、相関値は、さらなる洞察及び有効性を提供する物理量に対応しない。こうして、相関値は、かなりノイズの多いデータでは容認できる判定基準とはならない。

20

【0006】

その結果、ノイズを有するデータセットを処理し、そして当該データセットが統計的に線形であるかを同定するシステム及び方法であって、上記及び他の問題を克服する方法を提供することが望まれている。

【発明の開示】

【0007】

従って、本発明の実施態様は、データが統計的に線形挙動を示すかを決定するためにデータを処理する方法及びシステムを提供する。統計的線形データは、当該データが一般的に下方向又は上方向に曲がらず、又はそうでない場合に当該データが増幅を示さないということを意味する。このようなデータは、一般的に、当該データ上に重ねられた大きいノイズシグナルを伴ってもおおまかに線形であると表される。当該データは、リアルタイムPCR処理又は増幅若しくは増殖を示す他の過程から得ることができる。

30

【0008】

本発明の代表的な実施態様では、データの処理方法が提供される。当該方法は、一般的にシグナル成分とノイズ成分とを有するデータポイントのオリジナルセットを得ることを含む。オリジナルデータセットは線形関数に近似される。1の態様では、近似は、当該データセットに対して最小二乗法近似を計算することにより達成される。当該方法は、オリジナルデータセットと線形近似との間の残差を計算し、そしてオリジナルデータセットと線形近似との間の残差の尺度(measure)を計算することを含む。1の態様では、当該尺度は標準偏差である。

40

【0009】

当該方法は、一般的に平滑化データセットを計算し、そして平滑化データセットとオリジナルのデータセットとの間の残差を計算することにより、データセット中に存在するノイズ成分を見積もることを含む。平滑化データポイントは、平滑化データポイントの近くに位置するオリジナルデータポイントの数値(values of original data points that are local to that smoothed data point)に基づく。1の態様では、平滑化データセットを計算するためにローパスフィルターが使用される。代表的なローパスフィルターは、Savitzky-Golayフィルター、デジタルフィルター、又はデジタル平滑化多項式フィルターを含む。別の態様では、平滑化データポイントの数値は、平滑化データポイントの周囲のウイ

50

ンドウ内に存在するオリジナルデータポイントの平均である。

【 0 0 1 0 】

当該方法は、一般的に、見積もられたノイズの残差の尺度を計算し、そして当該尺度を比較して当該オリジナルデータセットが統計的に線形挙動を示すかを決定することを含む。比較は、第一尺度と第二尺度の比を計算して、当該比が、所定の値より小さいか又は大きいかを決定することを含みうる。一の態様では、所定の値は1のオーダーである。

【 0 0 1 1 】

好ましい態様では、当該方法は独立型コンピュータ、ネットワーク接続コンピュータ、又はリアルタイムPCR装置などのデータ取得装置などのプロセッサにおいて実行される。リアルタイムPCR装置の一例は、BioRad Laboratoriesにより販売されるiCycler iQ Systemである。

10

【 0 0 1 2 】

本発明の別の代表的な実施態様では、当該曲線が線形挙動を示すかを決定するために、情報処理装置がデータ処理を実行するように命令するために適用された複数の指示を有する情報貯蔵媒体が提供される。1の態様では、情報貯蔵媒体は、RAM又はROMユニット、ハードドライブ、CD、DVD、又は他の持ち運び可能な媒体である。

【 0 0 1 3 】

本発明の別の具体的な実施態様では、PCR検出システムが提供される。PCR検出システムは、シグナル成分とノイズ成分とを有するデータポイントのオリジナルセットをもたらし検出器を含み、そして当該データが線形挙動を示すかを決定するためにデータを処理するロジックを含む。

20

【 0 0 1 4 】

図面及び特許請求の範囲を含む当該明細書の残りの部分を参照すれば、本発明の他の特徴及び利点が認識されよう。本発明のさらなる特徴及び利点、並びに本発明の様々な実施態様の構造及び実施は、添付の図面に関して以下に詳細に記載されている。図面において、同様の参照番号が同一又は機能的に類似する要素を示す。

【 発明を実施するための最良の形態 】

【 0 0 1 5 】

本発明は、データセットを処理し、そして当該データセットが統計的に線形であるかを同定し、並びにこのような線形データセットを増幅シグナルを含むデータセットから区別する技術を提供する。好ましい態様では、本発明は、PCR増幅又は増殖過程から得たデータを処理して、当該データをさらに解析する前に統計的線形データを同定し、そして取り除くために特に有用である。しかしながら、本発明の教示は、ノイズを含みうる全てのデータセット又は曲線、並びに特にノイズを含まなければ増殖又は増幅を示す曲線、例えば細菌増殖過程、に適用できるということが認められるべきである。

30

【 0 0 1 6 】

図1は、PCR曲線100の例を示す。ここで強度値110対サイクル数120が一般的なPCRプロセスについてプロットされている。強度値110は、目的の物理量のいずれかであり、そしてサイクル数は、当該処理の時間又は回数に関連する任意の単位であってもよい。このような増幅曲線は通常、図1に示されるように直線領域130の後に増幅領域140を含み、次に漸近領域150を有する。下向き曲線データなどのさらなる挙動タイプが存在することもある。増幅領域は、指数関数的、シグモイド様、高次多項式、又は他のタイプのロジスティック関数又はロジスティック曲線であって、増殖をモデル化する曲線を有してもよい。

40

【 0 0 1 7 】

関与する実験的プロセスを理解するために、増幅領域140の位置及び形を同定することは重要である。例えば、PCR過程において、基線領域(線形領域130)の末端で生じる増幅の開始を同定することが望ましいこともある。当該位置を同定するステップは、PCR過程が増幅を示さない場合、潜在的な増幅領域が存在するかを同定することに等しい。しかしながら、リアルタイムPCRデータがノイズを有するので、データセットが増幅

50

を示すかという同定、又は同等に、統計的に線形ではないという同定は難しいこともある。

#### 【 0 0 1 8 】

例えば、図 2 は、シグナル及びノイズを含むデータポイント 2 4 0 を有するデータセットから作成されたリアルタイム P C R 曲線 2 0 0 の線形領域 2 3 0 を示す。一定シグナルを產生する装置についてさえ、当該データを分析用のデータポイントに分けなければならない点に注意すべきである。ノイズは、データポイントにおける揺らぎを引き起こす。全体として、データは大まかに線形様式で上向きに移動する(つまりプラスの傾き)。しかしながら、曲線 2 0 0 は、点から点においてかなり非線形であるので、当該曲線にそった任意の 1 点で曲線 2 0 0 を直接分析することによっては全体として線形挙動は測定できない。曲線 2 0 0 の直接的な分析は、データが統計的に線形な挙動を示さないということを誤って決定するであろう。本発明の実施態様は、データが統計的に線形挙動を示すかを効果的に決定する。

10

#### 【 0 0 1 9 】

さらに、線形挙動を有するデータ曲線と増幅挙動を有するデータ曲線とを区別することは重要である。図 3 は、増幅を示すリアルタイム P C R 曲線 3 0 0 を示す。最初に、データは領域 3 3 0 で線形挙動を示し、そして後のサイクルにおいて領域 3 4 0 では増幅を示す。本発明の実施態様は、線形挙動のみを有する P C R 曲線 2 0 0 と、P C R 曲線 3 0 0 などの増幅曲線を有する可能性のある P C R 曲線とを確実にかつ一貫した正確性で区別する。

20

#### 【 0 0 2 0 】

図 4 は、本発明の実施態様に従って当該データが統計的線形挙動を示すかを決定するためにデータを処理する方法 4 0 0 を示す。当該データセットは、データポイントから構成され、そしてシグナル成分とノイズ成分を有する曲線を示す。

#### 【 0 0 2 1 】

ステップ 4 0 5 において、まずデータセットを集めるか又は受け取る。当該データセットは、多くのメカニズムを介して受け取られてもよい。例えば、当該データセットは、i C y c l e r i Q 装置又は類似する P C R 分析装置などの P C R データ獲得装置内にある(指示を実行する)プロセッサにより得られうる。当該データセットは、データが回収されるとすぐにプロセッサに供されてもよいし、又はデータセットはメモリーユニット又はバッファ中に格納され、そして実験が完了した後にプロセッサに供されてもよい。同様に、データセットは、ネットワーク接続(例えば L A N、V P N、イントラネット、インターネットなど)を介して、又は取得装置への直接的な接続(例えば、U S B 又は他の直接的な有線又は無線接続)を介してデスクトップコンピューターシステムなどの分離したシステムに提供されてもよいし、又は C D、D V D、又はフロッピーディスクなどの持ち運び可能な媒体で独立型コンピューターシステムに提供されてもよい。データセットが受け取られるか又は獲得された後に、当該データが分析されてもよい。

30

#### 【 0 0 2 2 】

ステップ 4 1 0 において、データセットへの線形近似が計算される。一般的に、近似はデータセットと近似との間の一致又は差異を計測するメリット関数 S を定義し、ここで低い数値のメリット関数は、典型的に当該近似についてより優れたパラメーターを示す。例えば、線形最小二乗近似では、メリット関数は、データ値 Y と近似関数  $f(x_i)$  との間の差の二乗であり、ここで N のデータポイントについて、以下の：

40

#### 【 数 1 】

$$S \equiv \sum_{i=1}^N (Y_i - f(x_i))^2$$

50

の通りである。PCR過程において、Yはデータ強度であり、そしてxはサイクル数である。図5Aは、PCR曲線200の線形近似510を示す。図5Bは、PCR曲線300の線形近似550を示す。

#### 【0023】

メリット関数は、異なるデータポイントについてメリット関数に対する異なる荷重貢献度(weight contribution)又は規格化因子(normalization factor)を含んでもよい。メリット関数は、データポイント値を測ってもよいし、又は差分を得る前にデータポイントの関数を取得してもよい。当該差は、1のx値におけるデータと異なるx値における $f(x)$ との間で取られてもよい。例えば、メリット関数の項(term)は、線形近似に直交する線のデータ曲線から線形近似までの長さを示すこともある。これは、線形近似が0の傾きを有さない限り、異なるサイクル数で生じる。当業者は、使用されうる多くの異なるメリット関数を認識するであろう。

10

#### 【0024】

ステップ415では、データと線形近似との間の残差Rが計算される。残差Rは、線形近似からのデータポイントまでの誤差に一致する値のセットである。例えば、残差は、線形近似値と各サイクル数における実際のデータポイントとの差であってもよく、 $R_i = Y_i - f(x_i)$ を与える。当該式は、残差の標準形態である。幾つかの実施態様では、残差は、線形近似のメリット関数を決定するために使用される数値に関する。他の実施態様では、残差は異なる数値である。図5Aでは、誤差520は、曲線200と線形近似510との間の残差Rの値を計算するために使用される。図5Bでは、誤差560は、曲線300と線形近似550との間の残差Rの値を計算するために使用される。

20

#### 【0025】

ステップ420では、データと線形近似との間の残差の尺度(measure) $\sigma_1$ が計算される。尺度は、残差の数値セットから作成される単一値である。1の実施態様では、残差は、標準偏差であり、以下の：

#### 【数2】

$$\sigma_1 = \sqrt{\frac{1}{N} \sum_{i=1}^N R_i^2}$$

30

が与えられる。幾つかの実施態様は、残差の各値について荷重値(weighting value)を有し、そして他の実施態様は、各残差の値又は全ての残差の値を、更なる又は他の関数にかけてもよい。当業者は、使用されうる多くの異なる尺度を認識するであろう。

#### 【0026】

ステップ425において、データセットに存在する見積りのノイズ成分が計算される。データは、2個の成分、真のシグナルとノイズからなると仮定される。こうして、ノイズは、真のシグナルと実際のデータポイントとの間の差異である。しかしながら、シグナルが検出される場合に、常にノイズが加わっているか又は存在しているので、真のシグナルは決して直接計測することはできない。

40

#### 【0027】

真のシグナルは、平滑化データポイントから構成される平滑化データセットとして見積もられる。図6Aは、PCR曲線300の平滑化データセット670を示す。平滑化データポイントの値は、当該平滑化データポイントの近くに位置する複数のオリジナルデータポイントの関数Gに基づく。近いという用語は、データポイントのx値が、計算されたデータポイントからどれだけ離れているかに関する。例えば、予定されたサイクル数(ウィンドウ)だけ異なる場合、ポイントは他のポイントの近くに位置するとされる。3～5回のサイクルのウィンドウは適切であるが、他のウィンドウは10又は20サイクル、又はそれより多くが使用されてもよい。様々なウィンドウ値が使用されてもよく、つまり各平

50

滑化データポイントは、異なるウィンドウで計算されてもよい。さらに、例えば分数のデータポイントが挿入される場合、分数のサイクル数を有するウィンドウが使用されてもよい。ウィンドウは、データポイントの周囲で対称でなくてもよく、つまり当該データポイントの前の1のポイントと後の3個のポイントが使用されてもよい。ポイントは、x値の差異が使用されるトータルスケール、つまり全サイクル数、に達すると、近くに位置することはなくなる。

【0028】

1の実施態様では、関数Gは、移動平均又はローパスフィルターである。例えば、関数Gは、所定のサイクル数の範囲内のオリジナルデータポイントの平均、例えば中心平均値(centered mean)をとりうる。こうして、1の態様では、以下の：

10

【数3】

$$G(x_L) = \frac{1}{2K+1} \sum_{i=L-K}^{L+K} Y_i$$

[式中、Lは、計算された平滑化データポイントの指数であり、そしてKは使用されるウィンドウである]

である。

20

【0029】

また、ステップ425では、平滑化データとオリジナルデータとの間の残差が計算される。当該残差は、見積もられたノイズであると定義される。平滑化データとオリジナルデータとの間の残差は、オリジナルデータと線形近似との間の残差と同じ様式で定義されてもよく、又は当該残差は異なる様式で定義されてもよい。図6Bは、PCR曲線300に関して見積もられたノイズ成分680と平滑化データセット670を示す。シグナル670上にノイズ成分680を重ね合わせると、データ曲線300を与える。

【0030】

ステップ430では、平滑化データとオリジナルデータとの間の残差の尺度 $\sigma_2$ が計算される。 $\sigma_2$ 値は、固有のノイズの増幅の尺度として使用される。1の実施態様では $\sigma_2$ は標準偏差である。尺度 $\sigma_1$ 及び $\sigma_2$ は、類似の又は異なる様式で定義されてもよい。

30

【0031】

ステップ435では、第一尺度 $\sigma_1$ は、当該データセットが線形挙動を示すかを決定するために、第二尺度 $\sigma_2$ と比較される。1の実施態様では、 $\sigma_1$ と $\sigma_2$ の比が得られる。当該比が所定の数値よりも小さいか又は大きい場合、当該データは線形挙動を示すと決定される。例えば、 $\sigma_1 / \sigma_2$ が1のオーダーの数値、例えば1.5より小さい場合、当該データは線形であると決定される。同様に、 $\sigma_1 < c_0 * \sigma_2$ が使用されてもよい。当該表現は、データと線形近似との間の差の尺度が、当該データに存在する見積もられたノイズの尺度の定数倍より小さくなくてはならないということを示す。幾つかの実施態様では、 $c_0$ の数値は変化してもよい。

40

【0032】

定数 $c_0$ は、ノイズの定義、並びに他の数値の定義が固有ではないと言う事実に関する。 $c_0$ の値は、多数のデータセットを試験することにより取得して、当該数字について理に適った数値を得ることができる。標準残差の標準偏差が使用される場合、定数( $c_0$ )について1.5の数値が上手く機能するということが試験により示唆された。他の残差及び残差の尺度が使用される場合、他の数値がより適することもある。一般的に、1のオーダーの $c_0$ の値は上手く機能するべきである。

【0033】

データが統計的にフラット(線形)である、例えば上向き又は下向きに曲がっていないか、又はそうでない場合に増幅を示していない、と判定されると、データはさらに分析をす

50



ることなく捨てられてもよい。

【 0 0 3 4 】

ある態様では、本発明のデータ処理技術を実行するためにプロセッサを制御するコード及び指示は、コンピュータが読み取り可能である情報貯蔵媒体、例えばRAM又はROMユニット、ハードドライブ、CD、DVD、又は持ち運び可能な他の媒体に格納される。

【 0 0 3 5 】

図7は、本発明の1の実施態様に従ったシステム700を記載する。図示されるように当該システムは、サンプル705、例えば細菌又はDNAを、サンプルホルダー710内に含む。当該サンプルからの物理的特徴715、例えば蛍光強度の値などは、検出器720により検出される。ノイズ成分を含むシグナル725は、検出器720からロジックシステム730に送られる。シグナル725から得たデータは、ローカルメモリー735又は外部メモリー740、又は貯蔵デバイス745に貯蔵されてもよい。1の実施態様では、アナログ-デジタルコンバーターは、アナログシグナルをデジタル形態に変換する。

【 0 0 3 6 】

ロジックシステム730は、コンピューターシステム、ASIC、マイクロプロセッサなどであってもよいし、又は含んでもよい。ロジックシステムは、ディスプレイ(例えば、モニター、LEDディスプレイなど)及びユーザー入力装置(例えば、マウス、キーボード、ボタンなど)を含んでもよいし、又はこれらに接続されていてもよい。ロジックシステム730及び他の成分は、独立型又はネットワーク接続コンピューターシステムの一部であってもよいし、又はサーマルサイクラー装置に直接取り付けられるか、又は取り込まれてもよい。ロジックシステム730は、プロセッサ750において行われる最適化ソフトウェアを含んでもよい。

【 0 0 3 7 】

1の実施態様では、ロジックシステム730は、データをプロセッシングし、そして統計的にフラットなデータを同定するための指示を含む。当該指示は、好ましくはダウンロードされ、そしてメモリモジュール735、740、又は745(例えば、ハードドライブ又は他のメモリー、例えばローカルRAM又はROM或いは取り付けられたRAM又はROM)に格納されるが、指示は、フロッピーディスク、CD、DVDなどの任意のソフトウェア貯蔵媒体上に提供されてもよい。本発明の実施態様ではコンピューターコードは、C、C++、Java(登録商標)、ビジュアルベーシックなどの様々なコード言語、或いはVBスクリプト、Java(登録商標)スクリプト又はPerlなどの任意のスクリプト言語、或いはXMLなどのマークアップ言語で実行されうる。さらに、様々な言語及びプロトコルは、本発明の態様に従って、外部及び内部記憶、及びデータ送達、及びコマンドにおいて使用することができる。

【 0 0 3 8 】

本明細書に記載されるプロセスが、例示的であり、そしてバリエーション及び変更が可能であることが認められよう。連続的に記載されたステップは、平行して行われてもよく、ステップの順番は変えられてもよく、そしてステップは変更されるか又は組み合わせられてもよい。

【 0 0 3 9 】

本発明が例示の方法により記載され、そして特定の実施態様について記載される一方、本発明が、開示された実施態様に限定されないということが理解されるべきである。反対に、本発明は当業者に明らかである様々な変更及び類似するアレンジに及ぶように意図される。その結果、添付の請求項の範囲は、その様な全ての変更及び類似するアレンジを包含するように最も広い解釈が認められるべきである。

【図面の簡単な説明】

【 0 0 4 0 】

【図1】図1は、PCR増幅曲線の例を示す。

【図2】図2は、ノイズ及び統計的に線形挙動を示すリアルタイムPCRデータセットを

10

20

30

40

50

示す。

【図 3】図 3 は、ノイズ及び増幅を示すリアルタイム P C R データセットを示す。

【図 4】図 4 は、本発明の実施態様に従って、当該データセットが統計的に線形挙動を示すかを決定するためにデータセットを処理する方法を示す。

【図 5】図 5 A は、統計的線形挙動を示すデータセットへの線形近似を示す。図 5 B は、増幅挙動を示すデータセットへの線形近似を示す。

【図 6】図 6 A は、本発明の実施態様に従ったリアルタイム P C R データの平滑化データセットを示す。図 6 B は本発明の実施態様に従ったリアルタイム P C R データの見積もられたノイズを示す。

【図 7】図 7 は、本発明の実施態様に従ったリアルタイム P C R データを処理するシステムを示す。

10

【図 1】

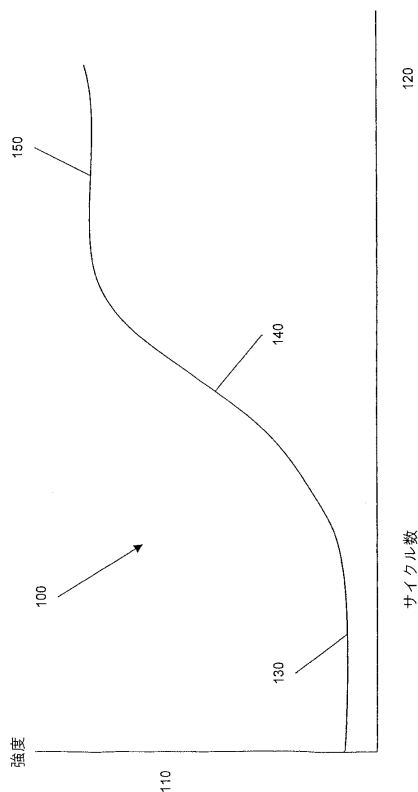


FIG. 1

【図 2】

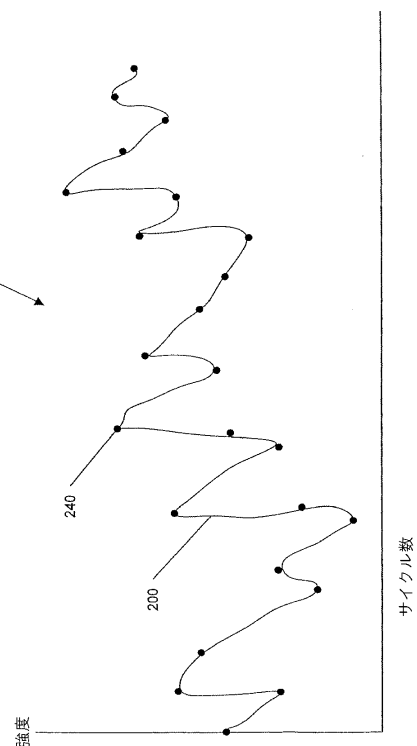


FIG. 2

【図 3】

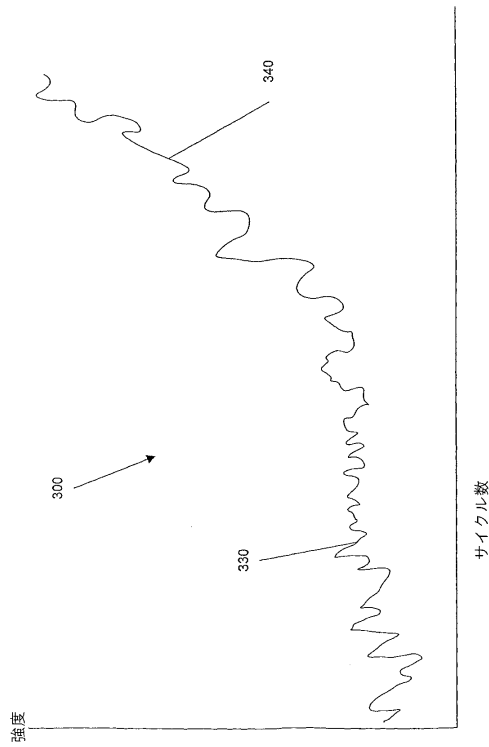


FIG. 3

【図 4】

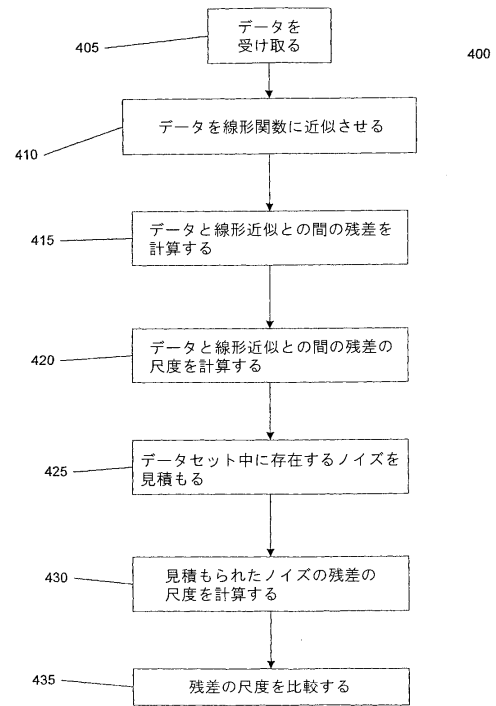


FIG. 4

【図 5 A】

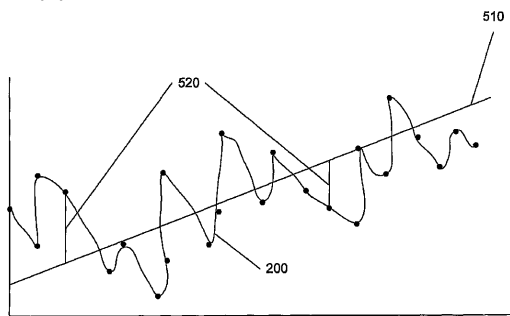


FIG. 5A

【図 6 A】

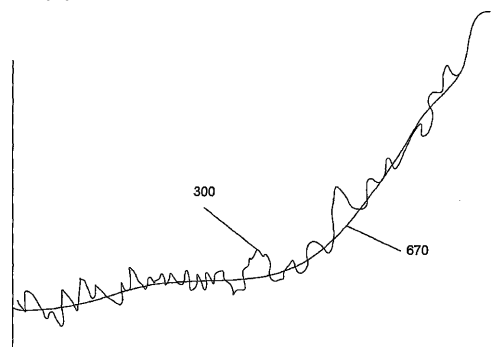


FIG. 6A

【図 5 B】

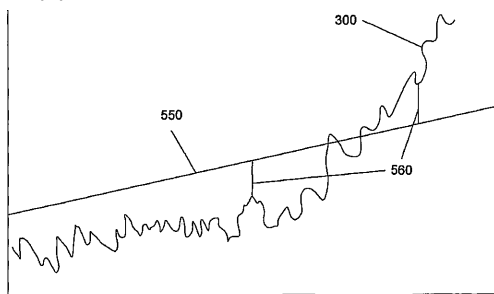


FIG. 5B

【図 6 B】

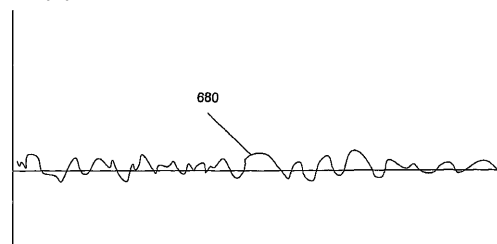


FIG. 6B

## 【図 7】

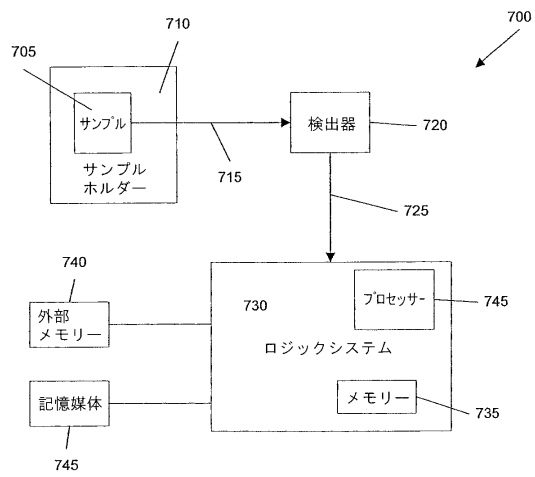


FIG.7

---

フロントページの続き

(74)代理人 100108903

弁理士 中村 和広

(72)発明者 ラーナー, ジェフリー

アメリカ合衆国, カリフォルニア 94606, オークランド, ナインス アベニュー 1004

審査官 吉田 美彦

(56)参考文献 国際公開第03/067215(WO, A1)

(58)調査した分野(Int.Cl., DB名)

G06F 17/18

C12M 1/00

C12Q 1/68